

Sample	Chimeras	Clone ID									
1R scrape	4	02e11	01h05	01g01	01g11						
1R swab	1	172g05									
1R punch	2	25h02	25f01								
1L scrape	9	05d11	05c05	05h08	06a02	06g07	06f04	06g09	161f01	05e10	
1L swab	6	09h06	10a05	09g12	09b09	10e08	10b03				
1L punch	7	12c09	08b12	08a08	12e10	12g10	11h11	08d09			
2R scrape	6	214f01	213f05	214a02	215c02	214g10	214a07				
2R swab	6	61h02	62b05	61f08	61c06	61c10	61b09				
2Rpunch	1	63a08									
2L scrape	7	220d09	217e01	217c12	219a04	219g09	219h12	218b09			
2L swab	2	222e03	223f08								
2L punch	5	72h11	67h03	68b07	71e04	68b04					
3R scrape	3	74e09	73g12	73a03							
3R swab	4	78g09	77a04	78e04	77d03						
3R punch	3	80f07	80a11	75c04							
3L scrape	5	81h08	81f11	82c03	81e05	81c09					
3L swab	5	227f09	227h09	228b05	225h12	228b03					
3L punch	5	84h10	83c12	87a06	84f11	83a03					
4R scrape	5	229c11	231b01	230d12	231h07	230g11					
4R swab	1	235h02									
4R punch	6	92d09	96c04	92h05	91f05	95f02	91b08				
4L scrape	6	98a02	97h08	97d02	97a07	97a09	98e12				
4L swab	3	101g02	102f12	102e09							
4L punch	1	237b11									
5R scrape	10	106a09	105e08	105g04	106c02	106h04	106f03	106h07	106c09	105h09	106c03
5R swab	1	242g05									
5R punch	3	107c01	111c02	112d08							
5L scrape	4	113d12	113f07	114c09	113f04						
5L swab	1	246c03									
5L punch	1	120c08									
mouse B1	5	176a01	175e03	176e02	176e08	175h09					
mouse B2	10	181b05	182f01	181c02	181a08	182g12	182b08	181a09	181g09	181g02	182e05
mouse B3	2	188b08	188e02								
mouse B4	0										

Table S5. Summary of chimeras removed before analysis. The clone ID numbers are listed and correlate with the GenBank record. Putative chimeras were deposited in GenBank and are annotated as such, as it is unclear if these sequences are true chimerics or rare OTUs. The Greengenes version of Bellerophon was used for chimera detection purposes (DeSantis et al 2006). This version differs from the original Bellerophon in that it calculates a divergence ratio, a measure of chimeric likelihood according to similarity to the parent sequences, to call chimeras. Additionally, it queries sequences against the Greengenes Core dataset in addition to the user supplied dataset. Settings used include: similarity to core set threshold=97; match length to core set threshold =1250; window size = 300; count of similar sequences to search for each window = 7; lanemask filter = on; parent to fragment threshold = 90; divergence ratio threshold = 1.1. Sequences with a divergence ratio >1.1 were removed from further analysis. A total of 140 sequences were removed after iterative queries following removal of chimeras.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. 2006a. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**: 5069-5072.